DOCUMENT RESUME

ED 277 737 TM 870 014

AUTHOR Yonker, Robert J.
TITLE High School and Beyond Cognitive Test: Evidence of Criterion-Related Validity and Test-Retest Reliability.
SPONS AGENCY Bowling Green State Univ., Ohio.
PUB DATE [Oct 86]
NOTE 17p.; This research was supported by a grant from the Faculty Research Committee.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Tests; College Entrance Examinations; *Concurrent Validity; Correlation; High Schools; *Longitudinal Studies; Standardized Tests; *Test Reliability; *Test Validity
IDENTIFIERS American College Testing Program; *High School and Beyond (NCES); Scholastic Aptitude Test

ABSTRACT

The High School and Beyond study is a national large-scale longitudinal survey of a representative sample of approximately 58,000 high school students. Information was collected at various stages from 1980-1986, on over 1,500 variables including achievement in six different areas measured by a 126-item cognitive test. Previous analyses of the psychometric properties of the cognitive test indicated minimal evidence of validity and acceptable internal consistency reliability. The purpose of this study was to collect additional evidence of the validity and reliability of the results of the cognitive test. Two year test-retest reliability coefficients were of a high magnitude with the exception of the civics subtest. Criterion-related validity evidence was investigated by correlating sophomore and senior year cognitive test results with senior year self-reported grades, class standing, and American College Testing Program/Scholastic Aptitude Test (ACT/SAT) standardized test scores. With the exception of the civics subtest, sophomore and senior year subtest scores correlated highly with senior year ACT/SAT standardized test scores and moderately well with senior year self-reported grades and class standing. (Author/JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277737

High School and Beyond Cognitive Test: Evidence of Criterion-Related Validity and Test-Retest Reliability

by

Robert J. Yonker

Bowling Green State University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Yonker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

This research was supported by a grant from the Faculty Research Committee at Bowling Green State University.

Running Head: Reliability and Validity of the HS&B Cognitive Test

TM 870 014

Abstract

The HS&B study is a national large scale longitudinal survey of a representative sample of approximately 58,000 high school students. The base year survey was conducted in the Spring of 1980 with subsequent follow-up studies in 1982, 84, and 86. Information was collected on over 1500 variables including achievement in six different areas measured by a 126 item cognitive test. Previous analyses of the psychometric properties of the cognitive test indicated minimal evidence of validity and acceptable internal consistency reliability. The purpose of this study was to collect additional evidence of the validity and reliability of the results of the cognitive test. Two year test-retest reliability coefficients were of a high magnitude (.73-.88) with the exception of the civics subtest (.33). Criterion-related validity evidence was investigated by correlating sophomore and senior year cognitive test results with senior year self-reported grades, class standing, and ACT/SAT standardized test scores. With the exception of the civics subtest, sophomore and senior year subtest scores correlated highly (.64-.83) with senior year ACT/SAT standardized test scores and moderately well (.35-.53) with senior year self-reported grades and class standing.

The High School and Beyond (HS&B) study is a national large scale longitudinal survey of high school students of the 1980's (National Opinion Research Center, 1980). HS&B was the second in a series of longitudinal studies sponsored by the National Center for Educational Statistics (NCES). HS&B is similar to NCES's first study, the National Longitudinal Study of the High School Class of 1972, which began in 1972 with follow-up studies in 1973, 74, 76 and 79 (Riccobono, Henderson, Burkheimer, Place, & Levinsohn, 1981). HS&B's primary purpose was to collect information on the educational and occupational plans and activities of young people as they passed through the American educational system and took on their adult roles. The base year survey was conducted in the Spring of 1980 with subsequent follow-up studies in 1982, 84, and 86. The design of the study included a stratified national probability sample of over 1,000 high schools. Approximately 30,000 sophomores and 28,000 seniors enrolled in public and private high schools across the nation participated in the base year survey. These samples represent the nation's 10th and 12th grade populations, which totaled about 3,800,000 sophomores and 3,000,000 seniors in more than 24,000 schools in the spring of 1980. Questionnaires and a cognitive test were administered to each student in the HS&B sample. The cognitive test measured both verbal and quantitative abilities. In addition, the sophomores' test included achievement measures in science, writing, and civics. The student questionnaire covered school experiences, activities, attitudes, plans, selected background characteristics, and language proficiency.

Other groups of respondents provided other types of information. The administrator in each selected school completed a questionnaire about the school; teachers in each school were asked to make comments on students in the sample; twins in the sample were identified and their counterpart twins were also surveyed; and a sample of parents of sophomores and seniors was surveyed primarily for information on the financing of higher education. The total survey effort thus provided a comprehensive data base for policy analysis in education and other areas of the social science. In fact, a recent review of the literature revealed that over 276 studies have utilized the HS&B data base (National Opinion Research Center, 1985). The HS&B data are being used to help set national and state educational policies. Justifications for positions related to tax credits for financing attendance at private schools, the revision of high school curriculum, the lengthening of the school day, merit pay for teachers, and competency-based testing to mention just a few issues have all been based in part on HS&B data.

An issue of concern does exist, however, about the accuracy of some of the collected data. Heyns' and Hilton's (1982) preliminary analyses of the psychometric properties of the cognitive test indicated minimal evidence of validity and acceptable internal consistency reliability except in the case of the civics subtest. The authors' conclusion was that an appropriate assessment of the technical characteristics of the cognitive test would not be possible until follow-up data became available. Given the current availability of the 1982 and 1984 follow-up data, the purpose of this study was to collect additional evidence of the validity and reliability of the

results of the cognitive test. Specifically, test-retest reliability evidence was investigated by correlating the cognitive test scores obtained in the sophomore year with those obtained in the senior year. Criterion-related validity evidence was investigated by correlating sophomore and senior year cognitive test results with senior year self-reported grades, class standing, and standardized test scores, specifically ACT and SAT.

METHOD

Subjects

The base year survey was conducted in the spring of 1980. Students were selected through a two-stage, stratified probability sample with schools within a stratum drawn with a probability proportional to their size. Once a school was selected, up to 36 sophomores and 36 seniors were drawn randomly from the students enrolled in each selected school. The total number of schools selected for the sample was 1,122 from a sampling frame of 24,725 schools with grades 10 or 12 or both. The sampling frame, defined as the universe of high schools in the United States, was obtained from the 1978 list of U.S. elementary and secondary schools of the Curriculum Information Center and supplemented by the National Center for Educational Statistics lists of public and private elementary and secondary schools. Of 1,122 possible schools, students at 1,015 schools completed survey questionnaires. Some schools having fewer than 36 sophomores or 36 seniors, student absences on the survey date, and declining to participate resulted in a completion rate of 84 percent: 30,030 sophomores and 28,240 seniors out of the 69,662 eligible students (Frankel, Kohnke, Buonanno, & Touvangeau, 1981).

The first follow-up was conducted in the spring of 1982. Of the 30,030 sophomores that participated in the base year survey approximately 95 percent completed follow-up questionnaires. A sample of over 12,000 of the 28,240 seniors of 1980 was followed-up primarily by a mailed survey. Completed questionnaires were obtained from 94 percent of those sampled (Jones, et al., 1983).

This study chose to investigate evidence of criterion-related validity of the HSB cognitive tests by correlating HS&B test scores with ACT and SAT test scores. ACT and SAT scores were available for only those sophomores included in the transcript survey. During the fall of 1982 high school transcripts were requested for a sample of the 29,737 seniors who participated in the survey in 1980 as sophomores. Of the 18,152 requests for transcript information, 15,941 were received (Jones, Knight, Butz, Crawford, & Stephenson, 1983). Thus, the subjects for this study were 1980 sophomores who were part of the transcript study that completed both the base year and first follow-up cognitive test. The actual number of subjects for which information was obtained varied because of missing data and is shown in Table 2.

Instrumentation

The data for this study were obtained from high school transcripts, student responses to the first follow-up questionnaire, and both base year and first follow-up test scores. The cognitive test administered to the sophomores in the spring of 1980 was readministered in the spring of 1982. The cognitive test utilizing multiple-choice item format covered the following areas:

1. vocabulary - A brief test using synonym format.

2. reading - A test based on short passages with several related questions concerning a variety of reading skills (analysis, interpretation) but focusing on straightforward comprehension.
3. mathematics - Quantitative comparisons in which the student indicated which of two quantities is greater, or asserts their equality, or the lack of sufficient data to determine which quantity is greater.
4. science - A brief test of science knowledge and scientific reasoning ability.
5. writing - A test of writing ability and knowledge of basic grammar.
6. civics - A test of knowledge of various principles of law, government, and social behavior.

Table 1 presents for each subtest the number of items, administration time, and internal consistency reliability estimates as reported by Rock, Hilton, Pollack, Ekstrom, and Goertz (1985).

Table 1

Characteristics of the HSB Cognitive Test

Subtest	Number of Items	Number of Minutes	Alpha Reliability
Vocabulary	21	7	.80
Reading	20	15	.77
Mathematics	38	21	.87
Science	20	10	.74
Writing	17	10	.80
Civics	10	5	.52
TOTAL	126	68	--

Data Collection

Base year data were collected between February 1 and May 15 of 1980; whereas, first follow-up data were collected between February 15 and June 11, 1982. One to two weeks prior to data collection an orientation was held in each school in order to inform sampled students about the objectives of the study and to brief them on the requirements of participation, the voluntary nature of the study, and the procedures for protecting the confidentiality of their responses. Representatives from the National Opinion Research Center assisted by the in-school designated HS&B coordinator administered the questionnaires and tests in the school during school hours to groups averaging approximately 20 students in number. Students were given one hour to complete the questionnaires followed by a ten-minute break. Immediately following the break the students took the six timed cognitive subtests. Data collection procedures may have deviated somewhat from the described process if the sample student transferred schools, quit school, or were absent; these procedures are described in detail elsewhere (Jones et al., 1983).

RESULTS

The descriptive statistics presented in Tables 2-4 were obtained by applying a weighting procedure to the unweighted sample. The weighting procedure was devised to take into account both differential selection probabilities for sample members and differential response rates for different types of schools and students. Jones et al. (1983) stated that the complex sample design of HS&B virtually ensures inaccurate results if the data are analyzed on an unweighted basis.

The means and standard deviations in Table 2 are those estimated for the population, the size of which is indicated in the third column.

Table 2

Sample and Estimated Population Size, Means, and Standard Deviations of Selected HS&B Variables

Year	Sample	Population	M	SD
Sophomore				
HSB Vocabulary	12,322	2,770,861	11.00	4.33
HSB Reading	12,322	2,772,799	9.23	3.88
HSB Mathematics	12,207	2,747,213	19.06	7.35
HSB Science	12,211	2,745,412	11.11	3.69
HSB Writing	12,097	2,715,221	10.50	3.90
HSB Civics	12,004	2,695,843	5.91	2.03
HSB Composite	11,880	2,673,707	67.11	21.04
Senior				
HSB Vocabulary	12,342	2,774,649	12.74	4.63
HSB Reading	12,322	2,756,621	10.20	4.18
HSB Mathematics	11,970	2,687,792	20.48	8.09
HSB Science	12,047	2,704,838	11.87	3.76
HSB Writing	11,970	2,685,236	11.76	3.92
HSB Civics	11,932	2,674,315	6.75	2.11
HSB Composite	11,756	2,635,824	74.15	22.68
Self-Reported Grades	11,013	2,521,147	NA	NA
Class Centile Standing	9,093	2,057,252	53.12	28.27
SAT Mathematics	2,887	543,796	465.74	112.99
SAT Verbal	2,857	541,654	431.35	107.83
ACT English	2,164	492,237	18.42	5.23
ACT Mathematics	2,158	491,064	17.66	7.75
ACT Social Science	2,163	492,380	17.91	7.36
ACT Natural Science	2,160	491,282	21.32	6.18
ACT Composite	2,160	492,666	18.95	5.77

The means and standard deviations for the HS&B subtests reported in Table 2 are in number correct units; whereas, the correlations presented in Tables 3-4 were calculated using standardized scores based on raw scores corrected for guessing. The mean and standard deviation for self-reported grades are not given in Table 2 because

the information was collected on an eight point scale; thus, producing only an ordinal scale of measurement. The median response, however, to the question "Which of the following best describes your grades in high school during the First Follow-up survey?" was "mostly B's."

The coefficients shown in Tables 3 and 4 were obtained via the Pearson product-moment correlational procedure. A one-tailed alpha level of .01 was used to test for statistical significance of the coefficients. The p-value, not shown, for each coefficient in Tables 3 and 4 was less than .01; therefore, each coefficient can be considered to be greater than zero and of the magnitude indicated by the value of the coefficient. The composite score used in the calculations is the arithmetic average of the six subtests' standardized scores.

Table 3
Correlations Among Sophomore and Senior Year HS&B Cognitive Subtests

	Sophomore Year Subtests						
Senior Year Subtests	Voc	Read	Math	Sci	Write	Civics	Comp
Vocabulary	.79[1]	.69	.62	.65	.63	.54	.79
Reading	.69	.74[1]	.64	.64	.63	.53	.78
Mathematics	.61	.63	.80[1]	.63	.59	.47	.76
Science	.64	.64	.62	.73[1]	.54	.48	.74
Writing	.59	.60	.59	.54	.73[1]	.50	.72
Civics	.37	.37	.34	.34	.36	.33[1]	.43
Composite	.77	.76	.75	.74	.72	.60	.88[1]

[1]The coefficients on the diagonal are two-year test-retest reliability coefficients.

The coefficients on the diagonal in Table 3 are two year test-retest reliability coefficients. With the exception of the civics subtest, all of the test-retest reliability coefficients are

high in magnitude and greater than the correlations between a subtest and a different subtest.

Table 4

Correlations Among Sophomore/Senior Year HS&B Cognitive Subtests and Senior Year Self-Reported Grades, Class Centile Standing, and ACT/SAT Standardized Test Scores

Year	Self-Reported Grades	Class Centile	SAT Math	SAT Verbal	ACT Eng	ACT Math	ACT SS	ACT NS	ACT Comp
Sophomore									
Vocabulary	.39	.42	.56	.77	.67	.53	.68	.61	.71
Reading	.43	.47	.61	.74	.67	.57	.70	.65	.74
Mathematics	.46	.51	.78	.63	.61	.74	.61	.63	.75
Science	.36	.40	.62	.66	.55	.57	.63	.65	.69
Writing	.47	.53	.51	.59	.69	.48	.53	.49	.61
Civics	.35	.39	.40	.50	.50	.39	.52	.44	.52
Composite	.50	.56	.73	.81	.78	.69	.77	.73	.85
Senior									
Vocabulary	.42	.45	.56	.77	.66	.51	.69	.61	.70
Reading	.45	.49	.60	.71	.64	.55	.70	.65	.73
Mathematics	.47	.53	.83	.65	.60	.77	.60	.65	.76
Science	.36	.39	.63	.65	.53	.58	.63	.66	.69
Writing	.47	.52	.51	.59	.65	.44	.47	.44	.57
Civics	.23	.29	.31	.39	.31	.25	.34	.31	.34
Composite	.50	.56	.73	.81	.75	.69	.76	.74	.84

The coefficients in Table 4 can be interpreted as evidence of criterion-related validity of the HS& B cognitive test results. The correlation coefficients' magnitude ranged from moderate to high. The HS&B subtests correlated higher with other test results, specifically SAT and ACT, than they did with measures derived from high school grades, specifically self-reported grades and class centile standing. In general, a given HS&B subtest correlated higher with a standardized subtest of similar content than it did with standardized subtests with

dissimilar content. With the exception of the mathematics subtest, the HS&B composite score correlated higher with each of the criterion measures than did any HS&B subtest with a criterion measure. On the average, the lowest correlations with the criterion measures were obtained for the civics subtest.

DISCUSSION

The conceptualization of the study and interpretation of the results were based on the following assumptions about the nature of validity and reliability:

1. Validity and reliability refer to the results obtained with a test and not the test itself.
2. Validity and reliability are a matter of degree; neither exist on an all or none basis.
3. Validity is specific to some particular use and reliability refers to a particular type of consistency.
4. Validity is a unitary concept based on various kinds of evidence. The three basic methods of accumulating evidence to support the validity of an interpretation (content, criterion-related, and construct) simply provide convenient categories for describing the evidence of validity (Gronlund, 1985).

The high magnitude (.73-.88) of the two year test-retest reliability coefficients obtained in this study indicated with the exception of the civics subtest (.33) that the results produced by the HS&B subtests and the composite were consistent from the sophomore to the senior year. This consistency does not mean that the test scores remained the same from the sophomore to the senior year; in fact, on

the average the scores increased as one might expect in a learning situation. What it does mean, however, is that students who scored the highest in the sophomore year tended to score the highest as seniors and those who scored the lowest as sophomores tended to score the lowest as seniors. The low test-retest reliability of the civics subtest found in this study is consistent with the low internal consistency reliability of the civics subtest reported by Heyns and Hilton (1982) and Rock et al. (1985).

The description of the historical development of the HS&B cognitive test found in Heyns and Hilton (1982) and Rock et al. (1985) provided evidence of validity of a content nature; whereas, the results of this study offer criterion-related evidence of validity. With the exception of the civics subtest, HS&B sophomore and senior year subtest scores correlated moderately well (.35-.53) with senior year self-reported grades and class standing based on actual grade point average. Sophomore and senior year subtest scores correlated highly (.64-.83) with senior year standardized subtest scores of similar content. The magnitude of these correlations are close in value to the correlations reported between ACT/SAT scores and standardized achievement test batteries that have more items and whose administration time is longer. The composite HS&B test score correlated more highly with self-reported grades, class standing, and ACT/SAT standardized test scores than did any of the subtests. The only exception to this statement is that the HS&B mathematics subtest score correlated more highly with ACT/SAT mathematics test scores than did the HS&B composite score.

In summary, the fact that the HS&B cognitive test results correlated with other typical measures of achievement that is grades and ACT/SAT test scores should increase HS&B data base users' confidence that the cognitive test administered to a representative sample of U.S. high school sophomores and seniors was measuring aspects of knowledge in the identified achievement areas; however, the evidence collected thus far is not definitive. Since the nature of validity is a matter of degree and does not exist on an all or none basis, the completion of additional studies such as the degree to which changes in test scores can be predicted by course work and grades in related subjects during the last two years in high school may further increase HS&B data base users' confidence in the validity of the results of the cognitive tests.

References

Frankel, M., Kohnke, L., Buonanno, D., & Tourangeau, R. (1981). Sample design report. Report to the National Center for Educational Statistics by the National Opinion Research Center, University of Chicago.

Gronlund, N. E. (1985). Measurement and evaluation in teaching (5th ed.). New York: Macmillan.

Heyns, B., & Hilton, T. (1982). The cognitive tests for High School and Beyond: An assessment. Sociology of Education, 55, 89-102.

Jones, C., Clark, M., Mooney, G., McWilliams, H., Crawford, I., Stephenson, B., & Tourangeau, R. (1983). High School and Beyond 1980 sophomore cohort first follow-up (1982): Data file user's manual. Report to the National Center for Educational Statistics by the National Opinion Research Center, University of Chicago.

Jones, C., Knight, S., Butz, M., Crawford, I., & Stephenson, B. (1983). High School and Beyond transcript survey (1982): Data file user's manual. Report to the National Center for Educational Statistics by the National Opinion Research Center, University of Chicago.

National Opinion Research Center. (1980). High School and Beyond: Information for users of base year (1980) data. Report to the National Center for Educational Statistics, University of Chicago.

National Opinion Research Center. (1985). Annotated bibliography of studies using data from High School and Beyond a national longitudinal study for the 1980's. Report to the National Center for Educational Statistics, University of Chicago.

Riccobono, J., Henderson, L., Burkheimer, G., Place, C., & Levinsohn, J. (1981). National longitudinal study: Base year (1972) through fourth follow-up (1979) data file user's manual. Report to the National Center for Educational Statistics by the Research Triangle Institute.

Rock, D., Hilton, T., Pollack, J., Ekstrom, R., & Goertz, M. (1985). Psychometric analysis of the NLS and the High School and Beyond test batteries. Report to the National Center for Educational Statistics by Educational Testing Service.